

Uma Avaliação Experimental dos Impactos da Política de Qualificação Profissional no Brasil

Ana Maria Hermeto Camilo de Oliveira*, Eduardo Luiz Gonçalves Rios-Neto†

Sumário: 1. Introdução; 2. Metodologias de Avaliação de Políticas de Qualificação; 3. Análise de sobrevivência; 4. Conclusões; A. Organização dos Dados; B. Médias de Variáveis, segundo o Grupo do Experimento, Amostra Total e Desempregados; C. Modelos Paramétricos; D. Análise Paramétrica.

Palavras-chave: avaliação de políticas públicas; treinamento e alocação de mão-de-obra; análise de sobrevivência.

Códigos JEL: J08; J64.

O principal objetivo deste artigo é determinar em que medida o Plano Estadual de Qualificação Profissional (PEQ) em Minas Gerais tem servido como instrumento efetivo para a redução do nível e da duração do desemprego dos seus participantes. A análise se baseia no estudo experimental de uma amostra de egressos do treinamento na Região Metropolitana de Belo Horizonte, que constitui o grupo de tratamento, incorporando informações sobre a evolução de um grupo de controle, desde dezembro de 1996 até dezembro de 2000, em um banco de dados longitudinais com painéis retrospectivos de informações mensais, que permite a avaliação da duração do impacto do treinamento sobre o desemprego, que tende a ser cumulativo. Foram utilizados métodos de análise de sobrevivência não paramétricos, paramétricos e semi-paramétricos. Os resultados apontam que o impacto foi positivo para o total da amostra, reduzindo o número e a duração dos episódios de desemprego. No entanto, os efeitos variam com o grau de desfavorecimento no mercado de trabalho, definido pela condição de desemprego no momento do treinamento. Os impactos estimados foram melhores para os que não enfrentavam esta barreira no mercado de trabalho.

* Afiliação: CEDEPLAR/UFMG. Endereço: Rua Curitiba, 832 – sala 717. 30315-430, Belo Horizonte, MG. Fone: (31) 3279.9181. Fax: (31) 3201-3657. E-mail: ahermeto@cedeplar.ufmg.br

† Afiliação: CEDEPLAR/UFMG. Endereço: Rua Curitiba, 832 - 9º. 30315-430, Belo Horizonte, MG. Fone: (31) 3279.9100. Fax: (31) 3201-3657. E-mail: eduardo@cedeplar.ufmg.br



This paper aims to evaluate the employment impacts of the National Vocational Training Program (PLANFOR). The evaluation is based on the experimental data basis designed to evaluate the case of Minas Gerais state in 1996. The findings indicate positive employment impacts. The employment impacts were felt more strongly with the increased duration after the provision of treatment. The data came from an experimental survey conducted by CEDEPLAR between 1996 and 2000, and provide retrospective occupational data. The employment impact is evaluated through the application of event history analysis (non-parametric, parametric, and semi-parametric) to the longitudinal experimental data. Results indicated that the unemployment duration of the control's group spells is greater than the treatments one. This difference is greater as the spell's length increases. There is no impact when the analysis is restricted to individuals that were unemployed in the beginning of the experiment.

1. INTRODUÇÃO

O objetivo deste artigo é determinar em que medida a política de qualificação profissional no Brasil tem servido como instrumento efetivo para melhorias nas condições de inserção no mercado de trabalho, no que se refere à redução do desemprego. O artigo apresenta a metodologia e resultados da avaliação de impactos do programa público de treinamento conduzido na Região Metropolitana de Belo Horizonte no âmbito do Plano Nacional de Qualificação Profissional (PLANFOR), que tem como objetivo central melhorar as condições de empregabilidade das populações mais vulneráveis no Brasil, priorizando os desempregados e outras populações excluídas.¹ Recentemente, o interesse neste tipo de política de trabalho vem aumentando, alimentado por preocupações sobre o nível e duração do desemprego dos trabalhadores com qualificação limitada e sobre a crescente desigualdade de rendimentos. A literatura internacional relativa a avaliações de impactos dos programas de treinamento governamentais para os economicamente desfavorecidos vem se acumulando por mais de três décadas, tendo sido consolidada na última década.

Abordagens convencionais à avaliação de programas sociais estimam impactos médios dos programas. Entretanto, há várias questões interessantes relativas à distribuição de benefícios dos programas, que exigem o conhecimento da distribuição dos impactos ou alguns de seus aspectos. Neste artigo, a questão abordada é se os impactos do programa são positivos e suficientes para produzir efeitos sobre os níveis e durações do emprego dos seus participantes. Neste caso, a eficácia poderia ser definida como o resultado efetivamente obtido com as ações em termos de geração e manutenção do trabalho. Há um crescente consenso de que programas de treinamento providos pelo setor público podem melhorar as perspectivas de emprego dos economicamente desfavorecidos (LaLonde, 1995, Bloom, 1984). Estudos sugerem que os ganhos resultam usualmente mais de aumentos da duração do emprego do que de aumentos dos salários por hora ou das horas trabalhadas por semana para os que já estão empregados.

Neste artigo, a análise se baseia em uma avaliação do impacto do PLANFOR sobre o desemprego, a partir de um estudo experimental conduzido pelo CEDEPLAR entre 1996 e 2000, que acompanhou uma amostra de egressos do treinamento, que constitui o grupo de tratamento, incorporando informações sobre a evolução de uma amostra de comparação (grupo de controle), através de painéis retrospectivos de informações mensais com cerca de 1200 indivíduos acompanhados ao longo de quatro anos. A abordagem experimental de avaliação de programas de treinamento avança em relação às comparações tradicionais “antes-depois”, que atribuem todas as melhorias aos resultados relativos aos níveis pré-programa avaliado. Contudo, se há melhorias gerais devido a efeitos econômicos amplos ou a efeitos

¹O PLANFOR foi elaborado em fevereiro de 1996 e cerca de 70 mil treinandos foram beneficiados com os recursos de 1996.

de ciclo de vida, os estimadores dos impactos do programa estariam viesados. A chave para resolver este problema é construir o resultado contrafactual, que representa o que teria acontecido na ausência do programa. O grupo de controle experimental fornece este resultado, permitindo identificar o que teria ocorrido aos participantes de um programa de treinamento, se não tivessem dele participado, em termos de emprego e renda, eliminando fatores de ciclo de vida e da economia ampla. As mudanças “antes-depois” para o grupo de tratamento – participantes – são então comparadas àquelas para o grupo de controle – não participantes. Os membros do grupo de controle eram elegíveis, se candidataram e foram aceitos inicialmente no programa, antes de serem excluídos aleatoriamente, constituindo o grupo mais similar possível ao grupo de tratamento. Dado que o status de tratamento é independente de características observáveis e não observáveis, estes dados experimentais são valiosos para identificar o efeito do treinamento.² Tendo sido verificada a boa qualidade dos grupos de tratamento e controle, não foi necessária a investigação dos determinantes da participação no programa.

Além do impacto imediato, busca-se avaliar a duração deste impacto, que tende a ser cumulativo, com base em informações coletadas de forma longitudinal, com a utilização de parâmetros de eficácia experimental. A questão do prazo de avaliação do desempenho após o programa é crucial para a avaliação de possíveis vieses. Um curto prazo de acompanhamento após o treinamento tenderia a levar a uma avaliação inadequada. No Brasil há poucos estudos de avaliação de programas sociais, e este experimento forneceu um banco de dados longitudinais único. O banco de dados construído, com informações sobre o estado dos indivíduos em vários pontos no tempo, permitiu utilizar modelos longitudinais sofisticados para analisar as transições – mudanças de estados na força de trabalho – ao longo do tempo, dentro do arcabouço dos métodos de análise de sobrevivência. Os objetivos desta análise são estimar, interpretar e comparar funções de sobrevivência e de risco entre os grupos de tratamento e controle do nosso experimento. Neste caso, a sobrevivência se refere à duração no estado de desemprego³ e o risco se refere ao risco de se ocupar. Além disto, buscamos verificar a relação de variáveis explicativas, como raça, idade, nível de escolaridade, sexo, entre outras, e o tempo de sobrevivência, através de modelos de regressão.

Os principais resultados encontrados apontam que o impacto de participar do PLANFOR sobre a duração do desemprego foi para o total da amostra, reduzindo o número e a duração dos episódios de desemprego. No entanto, os efeitos variam com o grau de “desfavorecimento” no mercado de trabalho, definido por aspectos como a condição de desemprego no momento do treinamento. Os impactos estimados foram melhores para os que não enfrentavam esta barreira no mercado de trabalho. Para os desempregados no momento do treinamento, a participação no treinamento não tem um efeito positivo significativo em termos de geração e manutenção do trabalho.

Tendo em vista os objetivos propostos neste trabalho, em primeiro lugar são apresentadas metodologias para avaliação de programas de treinamento profissional e descrita brevemente a nossa implementação de uma metodologia experimental e o processo de amostragem. Em seguida, efetuamos a análise da eficácia do PLANFOR, utilizando métodos de análise de sobrevivência para comparar as taxas de risco de saída dos episódios de desemprego empíricas dos grupos de tratamento e controle, com o propósito de captar os efeitos do treinamento, que correspondem ao impacto de ser membro do grupo de tratamento sobre as taxas de emprego alterando a duração dos episódios de emprego e desemprego.

2. METODOLOGIAS DE AVALIAÇÃO DE POLÍTICAS DE QUALIFICAÇÃO

Programas de treinamento e qualificação profissional podem ser definidos como uma tentativa de adicionar capital humano à força de trabalho. Embora estes programas não gerem empregos direta-

²Vale ressaltar que os resultados são condicionados ao fato dos indivíduos terem se inscrito no programa. Se de um lado isto torna possível comparar os dois grupos, de outra a generalização dos resultados para outros grupos não é automática.

³Desemprego aqui definido no sentido de desocupação, ou seja, incluindo os indivíduos que não trabalharam e que procuraram ou não trabalho no mês de referência.



mente, reduzem o desemprego na medida em que, ao expandir a qualificação dos trabalhadores, dá acesso aos desempregados a postos de trabalho vagos que não seriam acessíveis na ausência do treinamento. Dentre as variáveis dependentes que podem ser utilizadas no processo de avaliação do impacto de um programa de qualificação profissional, há duas que se destacam: a renda média e o emprego gerado depois do curso de treinamento. Nesse artigo, a ênfase é no emprego. Ao controlar por características individuais e correlacionar a probabilidade de emprego com os programas, deve-se ter em mente o contexto econômico e institucional. Deve-se levar em conta que a maioria dos empregos disponíveis para os grupos mais vulneráveis são no mercado secundário, onde as condições de trabalho são piores e os efeitos de longo prazo devem ser mais sistematicamente estudados, porque podem ser diferentes dos de curto prazo. Uma questão importante é avaliar se os empregos conseguidos são estáveis ou meramente oferecem um alívio temporário ao desemprego; assim, o problema dos grupos com dificuldades de colocação se torna um problema de dificuldades de manutenção.

Três metodologias são comumente utilizadas no processo de avaliação dos resultados dos programas: “antes/depois”, “não-experimental” e “experimental”. A metodologia de avaliação antes/depois diz respeito à avaliação do desempenho dos casos de tratamento, antes e depois da participação no curso de treinamento. As outras metodologias diferem e avançam por constituir um grupo de controle para medir a eficácia do programa: o desempenho do grupo de tratamento, antes e depois, medido em termos de emprego, é cotejado com o desempenho obtido pelo grupo de controle.

A metodologia de avaliação antes/depois é de fácil implementação, mas pode induzir a erros de avaliação. Este critério consiste em se comparar os indicadores de eficácia do grupo de tratamento antes e depois do programa. Definindo-se $t = 0$ para período antes do programa, $t = 1$ para período depois do programa, Y_0 para emprego antes do programa, Y_1 para emprego depois do programa, $Y_1 - Y_0$ é o ganho obtido pelo programa. A maioria dos estudos utilizando esta metodologia tende a mostrar resultados de ganhos com o treinamento. No entanto, alguns problemas podem decorrer desta estimativa. Primeiro, a auto-seletividade, segundo a qual apenas aquelas pessoas com avaliação subjetiva de que ganharão com o programa se candidatarão. Desta forma, os ganhos verificados se referem a como o programa afeta as pessoas motivadas a participar. Outro problema diz respeito ao fato de que as variações antes e depois do programa estão sujeitas a flutuações econômicas de conjuntura e tendências, que são independentes do tratamento. Como não há um grupo de controle, estas flutuações e tendências são inadequadamente incorporadas como ganhos com o treinamento.

Dado que, além das características observáveis, a motivação individual (heterogeneidade não observável) também afeta o desempenho relativo de tratamento e controle no mercado de trabalho, constata-se que a comparação entre os grupos é complicada. A literatura sobre avaliação sugere dois tipos de métodos que utilizam grupos de controle: não-experimental e experimental. A principal diferença entre estes métodos se refere à natureza do grupo de controle. No método não-experimental, o grupo de controle é composto de indivíduos que possuem características similares ao grupo de tratamento, utilizando uma fonte alternativa de informações.⁴ Este método utiliza a econometria para se corrigir o viés decorrente do problema de heterogeneidade não observada, já que o grupo de controle não é constituído através de um mecanismo de alocação aleatória sob controle do pesquisador. Os indicadores de eficácia são comparados entre os grupos e os resultados são interpretados em termos de “diferenças em diferenças”: diferenças entre os grupos das diferenças entre os indicadores nos períodos depois e antes do programa, que devem ser estatisticamente significantes. Todavia, uma comparação simples entre os indicadores de eficácia dos grupos não permite obter estimativas consistentes, porque os grupos podem não ser equilibrados em relação a todas as variáveis associadas ao resultado. Desse modo, uma diferença observada após o tratamento pode se dever ao efeito real do tratamento ou a diferenças pré-existentes entre os grupos. Como são as próprias pessoas que decidem se vão participar do treinamento, e assim ser parte do grupo de tratamento ou controle, não se elimina a dúvida sobre a existência de alguma

⁴Por exemplo, dados de pesquisas amplas sobre mercado de trabalho, como a Pesquisa Mensal de Emprego.

variável não-observada determinada previamente ao tratamento e que ocorre distintamente entre os grupos (Burtless, 1995).

Seguindo o critério experimental, o viés de auto-seletividade e heterogeneidade não observada não ocorre, uma vez que a alocação de indivíduos entre tratamento e controle é feita aleatoriamente. Ou seja, o método experimental consiste na condução de um experimento onde pessoas de um mesmo conjunto de candidatos ao programa de treinamento são alocadas aleatoriamente ao grupo de tratamento (no qual participam do treinamento) ou ao grupo de controle (no qual não participam do treinamento). Com isso, o viés de motivação é controlado e se garante que a alocação aos grupos é independente das características não observadas dos candidatos, as quais constituem a base do viés de seletividade. Após algum tempo, os grupos são comparados com relação a características supostamente influenciadas pelo tratamento e, caso existam diferenças estatisticamente significativas, devem ser causadas pelo efeito do tratamento já que, por meio da aleatorização, os dois grupos são basicamente idênticos. Neste caso, a comparação da diferença entre os desempenhos depois e antes entre os grupos de tratamento e controle é um indicativo do ganho causado pelo programa.

2.1. A Pesquisa Experimental com Participantes do PLANFOR na RMBH

Esta avaliação de eficácia do PLANFOR parte de um projeto de desenvolvimento metodológico, cuja grande novidade foi a efetivação de uma análise amostral e domiciliar de acompanhamento de participantes em quatro rodadas (*follow-ups*). O questionário âncora, ou primeira rodada, foi aplicado a uma amostra de 3721 entrevistados, sendo 2891 do grupo de tratamento e 746 do grupo de controle, em Dezembro de 1996. O primeiro *follow-up*, ou segunda rodada, foi a campo em Junho de 1997; e o segundo *follow-up*, ou terceira rodada, foi conduzido em Março de 1998 e conseguiu recuperar para Minas Gerais 2839 indivíduos (2212 do grupo de tratamento e 627 do grupo de controle). Uma decisão operacional tomada foi de restringir a amostra do terceiro *follow-up*, ou quarta rodada, realizada em Dezembro de 2000, somente para a RMBH, que contava com uma base de 1351 indivíduos. Foram realizadas 1159 entrevistas, ou seja, houve uma perda, após 33 meses decorridos da terceira rodada, de somente 14% da amostra. Portanto, para a RMBH, a série completa de dados retrospectivos remonta a Junho de 1996 e segue até Novembro de 2000.

Os procedimentos de amostragem e aleatorização do experimento adotados inicialmente tiveram de ser adaptados a diversas contingências encontradas ao longo do processo. O processo de amostragem consistiu de duas fases: a amostragem dos municípios representativos da implementação do programa em Minas Gerais e a amostragem do experimento aleatório com a análise dos cursos e candidatos e o sorteio daqueles que seriam treinados (tratamento) ou não seriam (controle). No ano de 1996 as limitações administrativas e a falta de informações acuradas dificultaram a condução de um experimento verdadeiramente experimental. Foi feita uma reformulação do plano amostral para adequar a amostra ao novo cenário surgido das investigações, tendo em vista a baixa demanda por alguns cursos. Apesar das dificuldades, várias delas comuns em pesquisas de amostragem, a coleta de dados trouxe as informações necessárias para a avaliação do programa de qualificação profissional em termos de sua eficácia no aumento da empregabilidade dos treinandos.

Uma análise da qualidade do processo de aleatorização no que tange às variáveis observáveis foi feita através de regressões logísticas para a determinação da participação no grupo de tratamento vs. controle para homens e mulheres na RMBH. A análise indicou que nenhuma covariável foi capaz de distinguir os indivíduos dos grupos de tratamento e controle; em outras palavras, o experimento parece ter sido bem sucedido (Rios-Neto e Assunção, 1998). A análise univariada mostrou uma pequena diferença de composição por sexo entre os grupos tratamento e controle. Parte desta diferença pode ser explicada pelo desenho amostral que não teve intenção de balancear estes fatores. Com relação às outras covariáveis (idade, nível de escolaridade, condição de atividade, rendimentos, sindicalização, participação anterior em treinamento), de maneira geral, os grupos tratamento e controle são equivalentes. A questão é se as covariáveis também são relevantes quando todas elas são consideradas simultaneamente; é

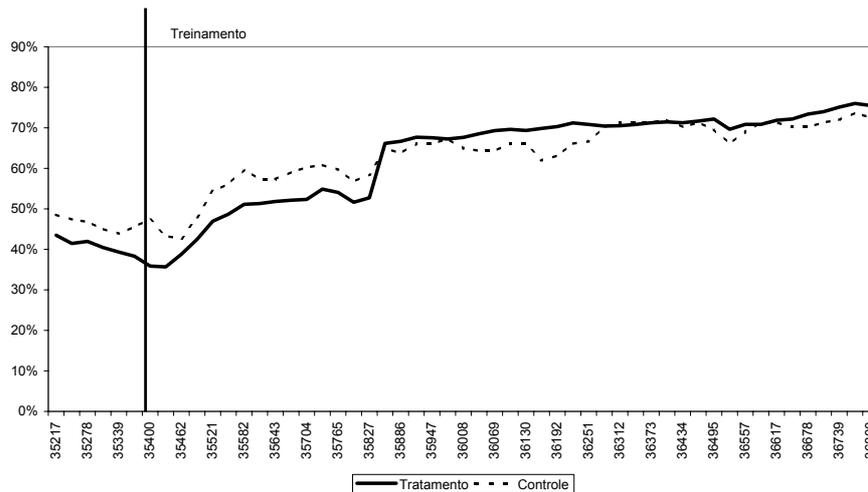


possível que variáveis julgadas importantes individualmente, perderem poder discriminatório entre os grupos quando outras são também consideradas. A análise multivariada, através de regressões logísticas,⁵ buscou responder esta questão, analisando as diferenças entre os grupos de tratamento e controle no que concerne às covariáveis observáveis de forma simultânea. Tanto para os homens, quanto para as mulheres, concluiu-se que as covariáveis não são capazes de diferenciar os grupos de tratamento e controle, e, portanto, a aleatorização parece ter funcionado muito bem na RMBH.

2.2. Análise Descritiva: Comparação entre os Grupos de Tratamento e Controle

Para analisar de forma descritiva possíveis variações no grupo de tratamento com relação ao emprego, tendo como referência o grupo de controle, que possam ser atribuíveis ao treinamento, verificamos se foram observadas mudanças nas taxas de ocupação do grupo de tratamento, em vários períodos após o curso de treinamento, em relação ao período anterior ao treinamento. Para esta análise, foram estimadas médias, mês a mês, das taxas de ocupação. As informações contidas no Gráfico 1 mostram a evolução das taxas de ocupação, ao longo do período pesquisado.

Figura 1 – Taxa de Ocupação, Segundo o Grupo do Experimento, RMBH, Jun/96-Out/00



Não verificamos de forma razoavelmente clara diferenciais entre os grupos de tratamento e controle ao longo do período pesquisado. Constata-se uma superioridade destas taxas para o grupo de controle no período anterior e no período imediatamente posterior ao treinamento. Após o período do treinamento, as taxas de ocupação de ambos os grupos assumem uma tendência ascendente, o que sugere a presença de efeitos de idade atuando sobre toda a amostra estudada. Após um médio prazo decorrido a partir do treinamento, as tendências dos grupos parecem mudar, o que se pode constatar através

⁵A variável dependente assume os valores 1, se o indivíduo é do grupo de tratamento e 0, se é controle. As covariáveis consideradas foram: idade, nível educacional, *dummies* para ser sindicalizado, ser associado, ter trabalhado na semana de referência, ter participado de treinamento nos últimos 5 anos. Sendo p a probabilidade de que um indivíduo pertença ao grupo de tratamento, o modelo de regressão logística trabalha com a chance relativa de pertencer ao tratamento vs. pertencer ao controle que, de acordo com o modelo, relaciona-se com as covariáveis através de uma equação linear no logaritmo da razão de chances de modo que $\log\left(\frac{p}{1-p}\right) = \alpha + \beta_1 X_1 + \dots + \beta_k X_k$. A interpretação do parâmetro β_i é dada em termos do efeito da covariável na razão de chances: se a covariável X_i sofrer um acréscimo de uma unidade, a razão de chances deverá mudar por um fator multiplicativo de e^{β_i} .

das diferentes inclinações das curvas. A série temporal de taxas de ocupação do grupo de tratamento assume uma inclinação maior, superando à curva do grupo de controle, o que indica uma maior ocupação de fato para os indivíduos que participaram de cursos de treinamento em relação aos que não participaram, a partir do médio prazo (ver Gráfico 1).

3. ANÁLISE DE SOBREVIVÊNCIA

3.1. Aspectos Conceituais da Análise de Sobrevivência Aplicada ao Estudo

Neste estudo, o plano de observações é misto: prospectivo e retrospectivo. Prospectivo porque o início da observação dos indivíduos é feito em um ponto bem definido no tempo e há um acompanhamento durante um período, com pesquisas em pontos posteriores no tempo. Tais pesquisas indagam aos indivíduos sobre as datas dos eventos, registrando através de quesitos retrospectivos os tempos (meses) nos quais o evento de interesse – transição do desemprego para a ocupação – ocorreu.

Para analisar os dados longitudinais obtidos, são utilizados de métodos estatísticos de análise de sobrevivência. O interesse principal é a ocorrência de eventos ou taxas de transições, e a acumulação de tempo em risco, ou duração, até (e incluindo) a ocorrência do evento ou censura. Neste estudo, o evento de interesse é a mudança qualitativa na força de trabalho, do desemprego para a ocupação; tal mudança é situada no tempo, correspondendo à transição de um estado discreto para outro: saída do estado de desempregado – entrada no estado de ocupado. Portanto, o estado de cada indivíduo na força de trabalho – desempregado ou ocupado – no mês da pesquisa é registrado e por quanto tempo este estado foi verificado, ou seja, a duração em cada estado. Tais dados são usados para modelar probabilidades de transição entre os estados. A unidade de análise não é o próprio evento, mas as trajetórias, ou seqüências individuais de eventos (e estados), ou duração (sobrevivência) nos estados de desemprego e ocupação.

A análise destas histórias de eventos, ou análise de sobrevivência, situa as transições individuais ao longo do tempo em seu contexto social, as relacionando com outras características. A questão é identificar como a participação no treinamento influencia a trajetória de cada indivíduo, mudando a probabilidade de transição do desemprego para o emprego, e como algumas de suas características o levam a se diferenciar dos outros indivíduos na medida que o tempo passa. Dado que diferentes seqüências de eventos não são igualmente prováveis, deseja-se estimar a distribuição de probabilidade destas trajetórias, que apresentam variáveis aleatórias de duração da permanência nos diferentes estados. Em outras palavras, a observação da história de eventos fornece um número de ocorrências de eventos, cuja distribuição ao longo do tempo deseja-se comparar entre os grupos do experimento.

Ao se atribuir um valor ao tempo de ocorrência do evento, implicitamente se está escolhendo uma escala e uma origem. A escala é a unidade comum na qual o tempo é mensurado: meses. Em estudos experimentais como este, a origem comum e bem definida corresponde ao tempo de aleatorização ao tratamento – dezembro de 1996 – após o qual os eventos estudados podem ocorrer. Como o principal objetivo é estimar o risco diferenciado associado ao treinamento, é somente no ponto de alocação ao tratamento que tais diferenciais de risco se tornam operativos. Portanto, a variável resposta de interesse é o tempo em meses a partir do início do acompanhamento dos indivíduos dos grupos de tratamento e de controle até a ocorrência do evento. A variável de tempo corresponde ao tempo de sobrevivência, porque indica o tempo que um indivíduo sobreviveu – permaneceu desempregado – ao longo dos meses de acompanhamento após o treinamento.

É importante ressaltar que eventos de transição para ocupação são repetidos, ou seja, podem ocorrer uma ou mais vezes com o mesmo indivíduo. As múltiplas ocorrências deste evento são observadas e registradas, sendo necessários métodos específicos de análise de sobrevivência para lidar com tais dados. Portanto, além das transições, a análise envolve as durações nos estados, ou seja, os indivíduos diferem no montante de tempo em risco, e esta exposição diferenciada tende a ser influenciada por vários fatores explicativos. A duração é uma variável contínua cujo valor é observado somente quando uma



transição ocorre; em geral, uma parcela da população não muda de estado dentro do período de observação ou é removida do risco por outras razões. Nestes casos, as durações são censuradas e os métodos para a análise de sobrevivência consideram este aspecto dos dados. A partir dos dados desagregados em episódios, as taxas de transição são modeladas como funções de variáveis independentes.

Diversos métodos de análise de sobrevivência, não paramétricos, paramétricos e semiparamétricos foram utilizados de forma complementar neste estudo. Os objetivos são: (1) estimar e interpretar funções de sobrevivência e risco, comparando-as entre os subgrupos da amostra: grupos do experimento – tratamento e controle; (2) estimar e comparar a relação entre variáveis explicativas e o tempo de sobrevivência dos diferentes grupos do experimento.

Nosso principal objetivo é encontrar a probabilidade do indivíduo permanecer desempregado, ou seja, sobreviver, após o tempo t , dado que estava desempregado até o tempo t . T corresponde à variável aleatória do tempo de sobrevivência do indivíduo, ou tempo de ocorrência do evento, cujos valores possíveis são não negativos ($T \geq 0$); t é qualquer valor específico de interesse para a variável aleatória T ; e δ é uma variável aleatória (0,1) que indica ocorrência do evento ($\delta = 1$) ou censura ($\delta = 0$). Todas as abordagens padrões para a análise de sobrevivência são probabilísticas; ou seja, os tempos nos quais os eventos ocorrem são assumidos como realizações de algum processo aleatório. Se segue que T é uma variável aleatória que tem uma distribuição de probabilidade. O que distingue diferentes modelos para dados de sobrevivência é a distribuição de probabilidade de T . As principais maneiras de descrever distribuições de probabilidade contínuas são as funções de sobrevivência e de risco.

A função de sobrevivência, $S(t)$, fornece a probabilidade de que o indivíduo permaneça desempregado além do tempo t , ou seja, indica a probabilidade de que a variável aleatória T exceda o tempo t : $S(t) = P(T > t)$, equivalendo ao produto das probabilidades condicionais de ter sobrevivido todos os prévios pontos do tempo. Sendo uma probabilidade, seus limites são 0 e 1. A função de risco (*hazard*), $h(t)$, é usada para descrever distribuições para dados contínuos de sobrevivência. Esta função é igual ao limite, na medida que Δt se aproxima de 0, de uma probabilidade relativa à sobrevivência, dividida por Δt , onde Δt denota um pequeno intervalo de tempo:

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t} \quad (1)$$

Sua interpretação conceitual é o risco instantâneo por unidade de tempo de um evento ocorrer no tempo t , dado que o indivíduo sobreviveu até este tempo. Isto significa que fornece o outro lado da informação dada pela função de sobrevivência: quanto maior $S(t)$ para um dado t , menor é $h(t)$. A probabilidade condicional fornece a probabilidade de que o evento ocorra no intervalo entre t e $t + \Delta t$, dado que o tempo de sobrevivência T é maior ou igual a t , isto é, dado que ele não tenha ocorrido até o início do intervalo. Para se ajustar pelo fato de que a probabilidade é uma função não decrescente de Δt – quanto maior o intervalo, mais provável é que um evento ocorra no intervalo – divide-se a probabilidade por Δt . Dada a condicionalidade, a função de risco é uma taxa condicional de ocorrência. O intervalo para esta razão varia entre 0 e ∞ , e depende da escala de medida do tempo.

3.2. Análise Não Paramétrica

Com o propósito de fazer uma investigação preliminar dos dados na forma de uma análise univariada sobre a duração até a ocorrência do evento, utilizamos métodos não paramétricos para estimar e comparar funções de sobrevivência. Foram gerados funções e gráficos de curvas de sobrevivência⁶ para a duração dos episódios de desemprego. No nosso desenho experimental, testes de hipóteses simples sobre diferenças da sobrevivência entre os grupos do experimento são relevantes para a análise dos dados; tendo sido o tratamento experimental aplicado a um grupo, mas não a outro, a questão é: “o

⁶Foram utilizados dois métodos: Kaplan-Meier e tábua de vida. Os resultados são similares e apresentamos aqui somente os referentes à tábua de vida, que agregam os períodos e assim são mais convenientes à nossa análise.

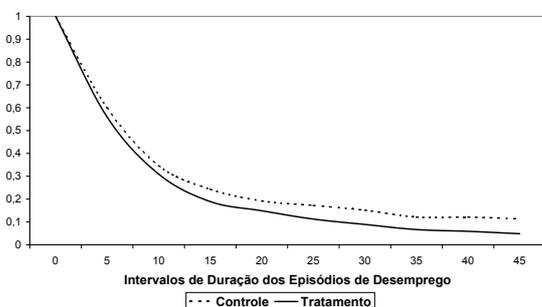


Figura 2 – Curvas de Sobrevivência da Amostra Total, segundo o Grupo do Experimento

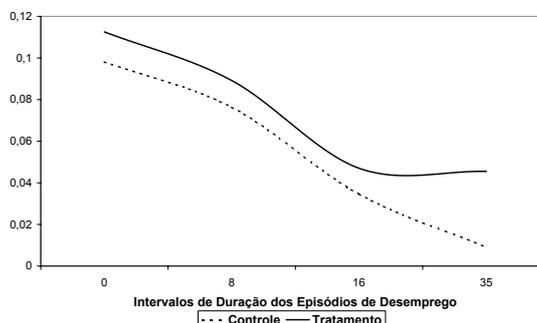


Figura 3 – Estimativas do Risco da Amostra Total, segundo o Grupo do Experimento

tratamento fez diferença na experiência de sobrevivência dos dois grupos?”. Dado que a função de sobrevivência fornece um relato completo da experiência de sobrevivência de cada grupo do experimento, uma função separada foi estimada para cada estrato corresponde ao grupo, sendo realizados testes de homogeneidade dos estratos. Neste caso, os métodos testam a hipótese nula de que as funções de sobrevivência são as mesmas nos grupos do experimento, ou seja: $S_1(t) = S_2(t)$ para todos os t , onde os subscritos distinguem os grupos. Tendo sido produzidas tabelas com de tábua de vida para cada grupo do experimento, foram feitos gráficos das funções de sobrevivência, para comparação das curvas. Além disto, foram especificadas covariáveis constantes no tempo e calculadas estatísticas lineares para testar as associações destas covariáveis e os tempos de ocorrência dos eventos, sem estimativas de parâmetros.

Os dados consistem de durações dos episódios de desemprego, mensuradas desde o mês de entrada no desemprego até o mês em que o desempregado se ocupou. As durações após o 47º mês são censuradas. Conhece-se o mês em que o desempregado se ocupou, tal que os tempos de sobrevivência têm valores 1 a 46 (ver anexo A).

Em primeiro lugar é analisada a sobrevivência para o total da amostra, comparando os grupos do experimento. O gráfico 2 mostra que as duas curvas de sobrevivência são bastante próximas nos curtos intervalos de duração dos episódios de desemprego. Estas curvas indicam que o grupo de controle sempre apresenta uma sobrevivência maior no estado de desemprego, ou seja, seus membros permanecem mais tempo desempregados e a diferença em relação ao grupo de tratamento tende a aumentar na medida em que aumenta a duração do desemprego, embora visualmente nunca seja muito grande.

O gráfico 3 mostra a função de risco de sair do estado de desemprego e se ocupar para os dois grupos do experimento. Os intervalos de duração dos episódios de desemprego foram definidos de acordo com uma proporcionalidade do número de casos. Assim, é possível verificar a importância da análise que vá além do curto prazo, considerando comportamentos distintos ao longo do tempo. O risco de sair do desemprego é sempre menor para o grupo de controle, mas as curvas seguem tendências diferentes. Enquanto para o grupo de controle, o risco é sempre descendente na medida em que aumenta a duração do desemprego, e parece até haver uma aceleração deste declínio no longo prazo, para o grupo de tratamento, o risco declina até um ponto de média duração dos episódios de desemprego, a partir do qual se estabiliza. Isto significa que o grupo de tratamento parece ser mais favorecido no sentido em que seu risco de sair do desemprego se mantém o mesmo após uma média ou longa duração do desemprego.

Em seguida, buscando o impacto específico do treinamento sobre os desempregados, estimamos a sobrevivência e o risco somente para a subamostra de desempregados no momento do treinamento. O gráfico 4 mostra que as curvas de sobrevivência dos dois grupos do experimento são quase indistinguíveis para os episódios de desemprego com duração mais curta. Assim como para o total da amostra,

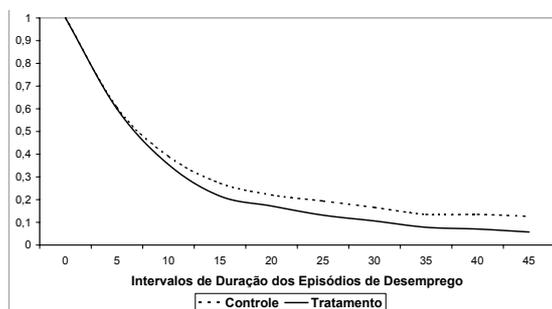


Figura 4 – Curva de Sobrevivência dos Desempregados, segundo o Grupo do Experimento

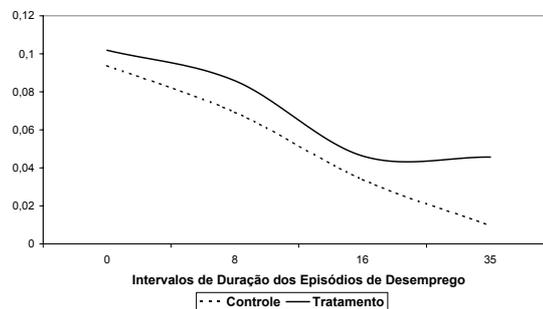


Figura 5 – Estimativas do Risco dos Desempregados, segundo o Grupo do Experimento

estas curvas também indicam que o grupo de controle tende a sobreviver mais no estado de desemprego, ou seja, seus membros permanecem mais tempo desempregados e a diferença em relação ao grupo de tratamento tende a aumentar na medida em que aumenta a duração do desemprego. Estas curvas da subamostra de desempregados no momento do treinamento são bastante similares às curvas da amostra total, mas a análise visual somente é insuficiente.

O gráfico 5 mostra a função de risco de sair do estado de desemprego e se ocupar para a subamostra de desempregados dos grupos do experimento. Os intervalos de duração dos episódios de desemprego foram definidos da mesma forma acima. O risco de sair do desemprego também é sempre menor para o grupo de controle e as curvas seguem tendências diferentes. Para o grupo de controle, o risco é sempre descendente na medida em que aumenta a duração do desemprego, e para o grupo de tratamento, o risco declina até o ponto de média duração dos episódios de desemprego, a partir do qual se estabiliza. Da mesma forma, isto significa que o grupo de tratamento parece ser mais favorecido no sentido em que seu risco de sair do desemprego se mantém o mesmo após uma média ou longa duração do desemprego. Este gráfico permite visualizar uma maior diferença da subamostra de desempregados em relação ao total da amostra. Os níveis de risco de saída do desemprego para os desempregados são sempre menores do que os riscos do total da amostra, o que sugere um impacto menor do treinamento para os desempregados.

A tabela 1, calculada a partir das estimativas de sobrevivência acima, confirma as diferenças entre os grupos do experimento, favorecendo o grupo de tratamento, o que, em princípio, sugere um impacto positivo do treinamento. A duração média do desemprego para o grupo de tratamento do total da amostra é de 10 meses e para o grupo de controle é de 12 meses. Para os desempregados, esta duração média é 1 mês maior para ambos os grupos do experimento. Tanto os gráficos quanto estes valores médios de sobrevivência não são suficientes para afirmarmos que a diferença entre os grupos do experimento é estatisticamente significativa.

Tabela 1 – Tempo Médio de Sobrevivência no Desemprego (em Meses), segundo o Grupo do Experimento, Total da Amostra e Subamostra de Desempregados

	Média		Nº de Observações	
	Tratam.	Controle	Tratam.	Controle
Total da Amostra	9.9	11.8	2096	332
Desempregados	10.9	12.7	1594	222

Testes da hipótese nula de que as curvas de sobrevivência são estatisticamente equivalentes para os dois grupos do experimento foram feitos através do cálculo de estatísticas alternativas: testes log-rank,

Wilkinson⁷ e estatística da razão de verossimilhança. Estas estatísticas se baseiam em testes χ^2 onde os valores observados de ocorrência de eventos são comparados com valores esperados se as distribuições fossem homogêneas, ou seja, são como uma função dos desvios dos números observados dos eventos dos números esperados. O p-valor do teste da razão de verossimilhança é menor do que os outros testes porque parte do pressuposto adicional de que a função de risco é constante em cada grupo, implicando em uma distribuição exponencial para os tempos de ocorrência dos eventos (tabela 2).

Tabela 2 – Estatísticas de Teste de Homogeneidade (p-valores) das Curvas de Sobrevivência dos Grupos do Experimento, Total da Amostra e Subamostra de Desempregados

	Total da Amostra	Desempregados
Log-Rank	0.0048	0.0281
Wilcoxon	0.0705	0.2737
-2Log(LR)	0.0003	0.0025

Segundo os resultados apresentados na tabela 2, para o total da amostra, os três testes são unânimes em rejeitar a hipótese nula de não haver diferença entre as curvas de sobrevivência dos grupos de tratamento e controle. Isto confirma a impressão visual e descritiva de que o treinamento tem um impacto positivo sobre a duração dos episódios de desemprego para o total da amostra. Buscando verificar se este impacto também pode ser inferido para o grupo de desempregados no momento do treinamento, os testes log-rank e razão de verossimilhança rejeitam a hipótese de não haver diferença entre os grupos do experimento, mas o teste Wilcoxon não a rejeita. Este teste Wilcoxon dá mais peso aos tempos iniciais do que os finais, sendo menos sensível a diferenças entre grupos que ocorrem nos tempos mais avançados. Neste sentido, é mais sensível para se medir diferenças no início da curva, ou a ausência delas, como pudemos visualizar no gráfico 4, no qual as curvas de sobrevivência dos desempregados nos dois grupos do experimento são quase indistinguíveis para os episódios de desemprego com duração mais curta. Novamente, estas constatações reforçam a importância da distinção entre prazos de durações do desemprego na avaliação dos impactos do programa de treinamento.⁸

3.3. Análise Semiparamétrica: Modelo de Riscos Proporcionais de Cox

Estimamos modelos de regressão de Cox, porque estes, ao contrário dos métodos paramétricos,⁹ não requerem a escolha de alguma distribuição específica de probabilidade para representar os tempos de sobrevivência, ou seja, é semi-paramétrico e, assim, mais robusto.

Assim como nos modelos de regressão, um objetivo da análise de sobrevivência é obter alguma medida do efeito que descreva a relação tratamento-resposta ajustada por variáveis relevantes exógenas, comparando a experiência de sobrevivência dos grupos. Na regressão linear, a medida do efeito é um coeficiente β , e, na logística, é uma *odds ratio* expressa em termos de uma exponencial de um ou mais coeficientes e^β . Na análise de sobrevivência, a medida do efeito é a razão de riscos (*hazard ratio*),

⁷O teste de Wilcoxon difere da estatística log-rank somente pela presença do número total em risco em cada ponto no tempo; assim é uma soma dos desvios dos números observados vs. esperados de eventos ponderada pelo total de expostos ao início de cada intervalo. Neste sentido, é mais sensível para se medir diferenças no início da curva.

⁸Dentro da análise não paramétrica, além dos testes de diferenças entre grupos, é possível testar se covariáveis quantitativas são associadas com o tempo de sobrevivência. Cada variável é testada individualmente, sem controlar pelas outras. Estas estatísticas são generalizações dos testes log-rank e Wilcoxon e podem ser interpretadas como testes não paramétricos dos coeficientes do modelo (ver anexo D).

⁹Fizemos estimativas de modelos paramétricos de regressão com dados de sobrevivência censurados usando o método de máxima verossimilhança, como uma extensão dos modelos não paramétricos para cobrir situações mais complexas. Como os resultados das análises paramétricas e semi-paramétricas são similares, apresentamos somente os resultados dos modelos de Cox e colocamos em anexo os resultados paramétricos.



expressa em termos de uma exponencial do coeficiente da regressão, e^β . Apesar de diferente da *odds ratio*, a razão de riscos tem uma interpretação similar da força do efeito: o valor 1 significa que não há efeito; o valor 10 significa que o grupo exposto tem dez vezes o risco do grupo não exposto. Quanto maior a probabilidade de sobrevivência no tempo t , menor a taxa de risco correspondente. Se a taxa de riscos para um grupo exposto no tempo t é maior do que para o grupo não exposto, a probabilidade de sobrevivência correspondente é menor. A distinção entre os métodos é o tipo de variável dependente usada: na análise de sobrevivência é a duração até o evento (e há dados censurados), na regressão linear é uma variável contínua e na logística é uma variável dicotômica.

Os modelos semi-paramétricos permitem um tipo de análise estratificada que é efetiva no controle de variáveis de confundimento e facilitam o ajuste por períodos de tempo nos quais um indivíduo não está em risco de um evento, acomodando mensurações diretas e contínuas de tempos de eventos. Cox (1972) inovou ao propor o modelo de riscos proporcionais e um método de estimação – verossimilhança parcial. O modelo básico é:

$$h_i(t, X) = \lambda_0(t) \cdot \exp\{\beta_1 x_{i1} + \dots + \beta_k x_{ik}\} \quad (2)$$

Esta equação diz que o risco para o indivíduo i no tempo t é o produto de dois fatores: uma função de risco de base (*baseline hazard*), $\lambda_0(t)$, que não é especificada, exceto que não pode ser negativa; e uma função linear de um conjunto de k covariáveis X_i fixas, as quais são exponenciadas. A função $\lambda_0(t)$ é vista como a função de risco para um indivíduo cujas covariáveis têm valor 0, ou seja, que prevaleceria caso X_i não existissem. Tomando o logaritmo em ambos os lados:

$$\log h_i(t) = \alpha(t) + \beta_1 x_{i1} + \dots + \beta_k x_{ik} \quad (3)$$

onde $\alpha(t) = \log(\lambda_0(t))$. Especificando $\alpha(t) = \alpha$, o modelo exponencial é obtido, e $\alpha(t) = \alpha * \log t$, o modelo Weibull é obtido. Todavia, a maior vantagem da regressão de Cox é que tais escolhas são desnecessárias e a função $\alpha(t)$ pode assumir qualquer forma, mesmo a de uma função em degraus. A função de verossimilhança para o modelo de riscos proporcionais pode, portanto, ser dividida em duas partes: 1) uma parte depende de $\lambda_0(t)$ e de β (vetor de coeficientes); 2) a outra parte depende somente de β . O que a verossimilhança parcial faz é descartar a primeira parte e tratar a segunda - a função de verossimilhança parcial – como se fosse uma função comum de verossimilhança. As estimativas são obtidas achando valores de β que maximizam a verossimilhança parcial, independente da forma da função de risco de base.¹⁰ β_i mede o efeito de cada variável X_i na função de risco.

O modelo de riscos proporcionais implica que o risco para um indivíduo i com características X^* é uma proporção fixa do risco para outro indivíduo j com características X . A razão de riscos (*hazard ratio*) para estes indivíduos então é:

$$\frac{h_i(t, X^*)}{h_j(t, X)} = \frac{\lambda_0(t) \exp\{\beta_1 x_{ik}^* + \dots + \beta_k x_{ik}^*\}}{\lambda_0(t) \exp\{\beta_1 x_{jk} + \dots + \beta_k x_{jk}\}} = \exp\{\beta_1 (x_{i1}^* - x_{j1}) + \dots + \beta_k (x_{ik}^* - x_{jk})\} = \theta \quad (4)$$

O que é importante nesta equação é que $\lambda_0(t)$ se cancela no denominador e no numerador, resultando em uma razão de riscos constante ao longo do tempo. Isto significa que a razão de riscos não depende do tempo, e, uma vez que X^* e X forem definidos, ela se torna constante. Esta propriedade dos riscos proporcionais, a constância da razão de riscos, implica que as curvas dos logaritmos dos riscos para quaisquer dois indivíduos são paralelas e não se cruzam.

¹⁰A função L de máxima verossimilhança é parcial porque se considera apenas as probabilidades dos casos de ocorrências do evento (falhas), não se considerando os casos censurados. A função L é o produto de vários L_j que representam a verossimilhança de 'falha' naquele momento, dado que tenha sobrevivido até aquele momento. Neste sentido os casos censurados são considerados como expostos até aquele momento.

A fórmula para a equação de verossimilhança parcial é válida somente para dados nos quais dois eventos não ocorrem ao mesmo tempo. Todavia, são comuns dados que contêm tempos de eventos empatados, sendo portanto necessária alternativas para lidar com estas situações. Utilizamos neste estudo um método exato que assume que há uma ordem verdadeira mas desconhecida para os tempos de eventos empatados. Isto porque assumimos que os eventos podem ocorrer em qualquer momento em um continuum de tempo, no caso ao longo de cada mês, sendo razoável supor que os empates são meramente o resultado de uma imprecisa mensuração do tempo, devido ao agrupamento dos dados contínuos, não empatados, e que há uma verdadeira ordenação no tempo para os eventos. Para estes dados de duração do desemprego, seria possível se ocupar em qualquer ponto no mês.

As variáveis utilizadas são: *trata* indica o grupo do experimento (tratamento=1 vs. controle=0); *ocupdez96* representa o estado na força de trabalho no momento do treinamento (1=ocupado vs. 0=desempregado); *sexo* (1=homens vs. 0=mulheres); *branca* (1=brancos vs. 0=outras raças); *jovem* (14-19 anos=1 vs. 20 anos ou + =0); *casado* (1=casado vs. 0=outros estados conjugais); *temfilho* (1=tem filhos vs. 0=não tem filhos); *estuda* (1=frequênta escola vs. 0=não frequênta escola); *serie4* (1=até 4 anos de estudo vs. 0=mais de 4 anos de estudo).¹¹

Foram então estimados os coeficientes e suas estatísticas associadas (tabela 3).¹² Foram realizados testes χ^2 Wald para a hipótese nula de que cada coeficiente é igual a 0 e o teste global da hipótese nula de que $\beta = 0$, isto é, todos os coeficientes são 0, usando as estatísticas de testes da razão de verossimilhança, score e Wald. Para a estatística da razão de verossimilhança, são mostrados os resultados para o modelo com covariáveis e para o modelo sem covariáveis. A estatística χ^2 da razão de verossimilhança é a diferença entre estes números. Tendo visto que são elevadas, os p-valores são pequenos, e conclui-se que ao menos um dos coeficientes das covariáveis não é 0 tanto para o total da amostra quanto para a subamostra dos desempregados.

Analisando os parâmetros estimados para cada covariável, a razão de riscos (*hazard ratio*), e^β para variáveis categóricas é interpretada como a razão do risco estimado para aqueles com valor de 1 sobre o risco estimado para aqueles com valor de 0, controlando pelas outras covariáveis. Por exemplo, para a variável indicativa do grupo do experimento, significativa para o total da amostra, o valor 1,227 significa que o risco de se ocupar para o grupo de tratamento é cerca de 23% maior do que o risco para o grupo de controle. Da mesma forma, este risco é maior para os ocupados no momento do treinamento e para os homens, e menor para os mais jovens, para os que têm filhos e para os menos escolarizados. Quanto aos desempregados, o efeito do treinamento sobre a duração dos episódios de desemprego não é estatisticamente significativo.¹³

Finalmente, estimamos modelos semi-paramétricos para os episódios ordenados de desemprego. Esta abordagem consiste em conduzir uma análise separada para cada episódio sucessivo, o que corresponde a estimar um modelo para o intervalo entre o início e o final do primeiro episódio, um segundo modelo para o intervalo entre o primeiro e o segundo episódios, e assim por diante. Isto é útil se esperamos que o modelo seja diferente de um episódio para o outro. De fato, como mostra a tabela 4, para o total da amostra, o efeito do treinamento só é significativo para os primeiros episódios de desemprego, no sentido de significar um maior risco de saída do desemprego; risco este mais expressivo para o segundo episódio. Do terceiro episódio para frente, ou seja, entre aqueles que recorrentemente se ocupam e se desempregam, o efeito do treinamento não é significativo.¹⁴ A tabela 5 reforça os mo-

¹¹No anexo B constam as médias destas variáveis para cada grupo na amostra total e na subamostra de desempregados.

¹²Não há uma estimativa de intercepto – aspecto característico da estimação de verossimilhança parcial; o intercepto é parte de $\alpha(t)$, a função arbitrária do tempo, que se cancela nas equações estimadoras.

¹³Foram feitas estimativas robustas levando em consideração efeitos aleatórios para testar a hipótese nula global e as hipóteses nulas de parâmetros individuais. Os resultados seguiram na mesma direção, sendo o efeito do tratamento significativo para o total da amostra, mas com uma razão de riscos maior (1,264); e não significativo para a subamostra dos desempregados.

¹⁴Do quarto episódio em diante, a hipótese nula global de que todos os coeficientes são iguais a zero é aceita.

**Tabela 3** – Estimativa dos Coeficientes e Significância Amostra Total e Subamostra de Desempregados

	Amostra Total		Desempregados	
	Estimativa do Parâm.	Razão de Riscos	Estimativa do Parâm.	Razão de Riscos
OCPDEZ96	0,3246*	1,384		
BRANCA	-0,0593	0,942	-0,0315	0,969
JOVEM	-0,1718**	0,842	-0,1949**	0,823
CASADO	0,0235	1,024	0,0286	1,029
TEMFILHO	-0,1912***	0,826	-0,3501*	0,705
ESTUDA	-0,0307	0,970	-0,0560	0,946
SERIE4	-0,1588*	0,853	-0,1554*	0,856
SEXO	0,6282*	1,874	0,6416*	1,900
TRATA	0,2044*	1,227	0,1162	1,123
Teste hip. Nula: $\beta = 0$	Rejeitada a 1%		Rejeitada a 1%	
-2 log L - sem covar.	12304,393		9330,206	
-2 log L - com covar.	12077,684		9186,210	

* significante ao nível de 1%;

** significante ao nível de 5%;

*** significante ao nível de 10%;

Nota: Número de Observações: Total da Amostra: 2427; Desempregados: 1816

delos anteriores, no sentido de que o efeito do treinamento para os desempregados não é significativo, independentemente da ordem do episódio de desemprego.

3.4. O Modelo Exponencial por Partes

Os modelos paramétricos assumem que o risco é uma função suave e relativamente simples do tempo. O modelo semiparamétrico de Cox é menos restritivo neste aspecto, mas não é capaz de testar hipóteses sobre a forma da função de risco. Uma maneira de obter alguma da flexibilidade do modelo de Cox, sem perder a capacidade de testar hipóteses, é empregar o modelo exponencial por partes (piecewise exponential), que exigiu uma reestruturação dos dados com registros múltiplos para cada indivíduo, dividindo a escala de tempo em intervalos e assumindo que o risco é constante em cada intervalo, mas pode variar entre os intervalos. Definindo uma série de J intervalos, com pontos de corte a_0, a_1, \dots, a_J , onde $a_0 = 0$ e $a_J = \infty$, o intervalo j é dado por $[a_{j-1}, a_j)$ e o risco para o indivíduo i teria a forma: $\log h_i(t) = \alpha_j + \beta x_i$, onde para $a_{j-1} \leq t < a_j$. Portanto, o intercepto pode variar de um intervalo para outro. Cada registro é tratado como uma observação distinta, com o tempo refixado em 0 no início de cada intervalo e as covariáveis fixas no tempo são replicadas em cada intervalo.

Para este estudo, dividimos o período de 47 meses de observações em quatro prazos: até 8 meses após o treinamento, entre 9 e 16 meses após o treinamento, entre 17 e 35 meses após o treinamento e acima de 36 meses após o treinamento. Esta divisão se baseou nos riscos estimados pelo modelo não paramétrico e parece ser bastante razoável para analisarmos as diferenças entre os riscos de sair do desemprego dos grupos do experimento, de acordo com a duração dos episódios de desemprego. O teste de Wald para as covariáveis individuais indicou que a variável referente aos prazos tem um efeito significativo, implicando que o risco não é constante ao longo do tempo de duração dos episódios de desemprego, mas pode ser constante dentro de cada prazo.

Tabela 4 – Estimativa dos Coeficientes para Episódios Ordenados da Amostra Total

	1º Episódio		2º Episódio		3º Episódio	
	Estimat. Parâm.	Razão Riscos	Estimat. Parâm.	Razão Riscos	Estimat. Parâm.	Razão Riscos
OCPDEZ96	0,4384 [*]	1,55	0,2815 [*]	1,325	0,0481	1,049
BRANCA	-0,1039	0,901	-0,0088	0,991	-0,057	0,945
JOVEM	-0,1544	0,857	-0,2869 ^{**}	0,751	-0,2334	0,792
CASADO	0,0274	1,028	-0,1809	0,835	0,1252	1,133
TEMFILHO	-0,1242	0,883	-0,0697	0,933	-0,5167 ^{***}	0,596
ESTUDA	-0,0876	0,916	0,0582	1,06	0,0306	1,031
SERIE4	-0,101	0,904	-0,2155 ^{**}	0,806	-0,2677 ^{***}	0,765
SEXO	0,6979 [*]	2,01	0,4882 [*]	1,629	0,6632 [*]	1,941
TRATA	0,1887 ^{**}	1,208	0,2418 ^{***}	1,274	0,2649	1,303
Hip.Nula $\beta = 0$	Rejeitada a 1%		Rejeitada a 1%		Rejeitada a 1%	
N.º Observ. Censuradas	1220		653		312	
	180 (14,8%)		113 (17,3%)		66 (21,2%)	

* significante ao nível de 1%;

** significante ao nível de 5%;

*** significante ao nível de 10%;

Tabela 5 – Estimativa dos Coeficientes para Episódios Ordenados da Subamostra de Desempregados

	1º Episódio		2º Episódio		3º Episódio	
	Estimat. Parâm.	Razão Riscos	Estimat. Parâm.	Razão Riscos	Estimat. Parâm.	Razão Riscos
BRANCA	-0,108	0,898	0,1226	1,13	-0,0502	0,951
JOVEM	-0,2096 ^{***}	0,811	-0,2901 ^{***}	0,748	-0,0782	0,925
CASADO	0,0957	1,1	-0,081	0,922	-0,0594	0,942
TEMFILHO	-0,3436 ^{***}	0,709	-0,2411	0,786	-0,4836	0,617
ESTUDA	-0,0849	0,919	0,013	1,013	-0,089	0,915
SERIE4	-0,1389 ^{***}	0,87	-0,1969 ^{***}	0,821	-0,1146	0,892
SEXO	0,7131 [*]	2,04	0,4771 [*]	1,611	0,7498 [*]	2,117
TRATA	0,1873	1,206	-0,0429	0,958	0,1723	1,188
Hip.Nula $\beta = 0$	Rejeitada a 1%		Rejeitada a 1%		Rejeitada a 1%	
N.º Observ. Censuradas	881		492		248	
	142 (16,1%)		91 (18,5%)		55 (22,2%)	

* significante ao nível de 1%;

** significante ao nível de 5%;

*** significante ao nível de 10%;



Tabela 6 – Estimativa dos Coeficientes do Modelo Exponencial por Partes e Significância para a Amostra Total e para a Subamostra de Desempregados

	Amostra Total	Desempregados
Intercepto	-3,0842*	-2,3813
OCPDEZ96	0,5825*	
BRANCA	-0,0154	-0,0351
JOVEM	-0,1364***	-0,1236
CASADO	0,0988	0,1005
TEMFILHO	-0,0899	-0,306***
ESTUDA	0,1701**	0,0544
SERIE4	-0,2502*	-0,1954*
SEXO	0,4227*	0,338*
TRATA	0,1475***	0,124
PRAZO 1	-0,8832*	-1,8185*
PRAZO 2	-1,069*	-1,3337*
PRAZO 3	-1,1241*	-1,5288*
Deviance	4477,624	2164,8254
Log-Verossimilh.	3590,812	-1972,4127
Multipl.Lagrange	67,9668*	1662,3428*

* significante ao nível de 1%;

** significante ao nível de 5%;

*** significante ao nível de 10%;

Amostra Total: Número de Observações: 4129, Valores

Censurados: 2777 Desempregados: Número de Observações:

2653, Valores Censurados: 1763

Sendo todas as covariáveis categóricas, reestimamos o modelo dentro do arcabouço dos modelos lineares generalizados para dados agrupados, tratando o número de eventos ocorrendo em uma célula de intervalo de tempo–covariável como tendo uma distribuição de Poisson. O parâmetro desta distribuição é tomado como uma função log-linear das covariáveis. A estatística de teste da log-verossimilhança, os coeficientes, o nível de significância, etc. estimados desta forma são idênticos aos estimados pelo modelo paramétrico, com os sinais invertidos na tabela. A tabela 6 mostra os resultados desta forma de estimação do modelo exponencial por partes aplicado aos dados da amostra total e da subamostra de desempregados.

Para o total da amostra, o padrão temporal estimado é monotônico: o risco estimado diminui do primeiro para o segundo prazos e do segundo para o terceiro. Além disto, a variável indicativa do grupo do experimento é estatisticamente significativa e seu coeficiente aponta para um risco de saída do desemprego 15% maior ao longo do tempo para o grupo de tratamento em relação ao grupo de controle. Ao contrário, para a subamostra dos desempregados, o padrão temporal estimado não é monotônico: o risco estimado aumenta do primeiro para o segundo prazos e diminui do segundo para o terceiro. Adicionalmente, a variável indicativa do grupo do experimento não é estatisticamente significativa. Todos estes resultados novamente apontam para a importância da análise particionada por prazos, que levam a conclusões diversas.

4. CONCLUSÕES

Este artigo apresentou a metodologia e os resultados da avaliação de impactos do programa público de treinamento conduzido na RMBH no âmbito do PLANFOR. Especificamente, procurou-se determinar em que medida este programa tem servido como instrumento efetivo para melhorias nas condições de inserção no mercado de trabalho, no que se refere à redução do nível e da duração do desemprego dos seus participantes.

O impacto de participar do PLANFOR sobre a duração do desemprego pareceu ser positivo para o total da amostra, reduzindo o número e a duração dos episódios de desemprego. No entanto, os efeitos variam com o grau de “desfavorecimento” no mercado de trabalho, definido por aspectos como a condição de desemprego no momento do treinamento. Os impactos estimados foram melhores para os que não enfrentavam esta barreira no mercado de trabalho. Para os desempregados no momento do treinamento, a participação no treinamento não tem um efeito positivo significativo em termos de geração e manutenção do trabalho. Um ponto fundamental a ser destacado é a questão do prazo de avaliação do desempenho após o programa; um curto prazo de acompanhamento após o treinamento tenderia a levar a uma avaliação inadequada. Dado o aumento do interesse neste tipo de política de trabalho, devido às crescentes preocupações sobre o nível e duração do desemprego dos trabalhadores com qualificação limitada, este artigo pretendeu contribuir para a discussão sobre metodologias de avaliação de políticas públicas no Brasil.

Referências Bibliográficas

- Alisson, P. D. (1995). *Survival analysis using the SAS System: a practical guide*. SAS Institute, Cary, NC.
- Allison, P. (1984). *Event history analysis*. Sage, Beverly Hills, CA.
- Ashenfelter, O. (1978). Estimating the effect of training program on earnings. *Review of Economics and Statistics*, 60.
- Bell, S., Orr, L., Blomquist, J., & Cain, G. (1995). *Program applicants as a comparison group in evaluating training programs*. Upjohn Institute for Employment Research, Michigan.
- Bivar, W. (1991). Estimativas da duração média do desemprego no Brasil. *Pesquisa e Planejamento Econômico*, 23:275–312.
- Bloom, D. E. (1984). On the nature and estimation of age, period and cohort effects in demographic data. Discussion Paper 84-7, Harvard University.
- Blossfeld, H., Hamerle, A., & Mayer, K. (1989). *Event history analysis: statistical theory and application in the social sciences*. Lawrence Erlbaum, Hillsdale, NJ.
- Burtless, G. (1995). The case for randomized field trials in economic and policy research. *Journal of Economic Perspectives*, 9(2):63–84.
- Card, D. & Sullivan, D. G. (1988). Measuring the effect of subsidized training programs on movements in and out of employment. *Econometrica*, 56(3):497–530. available at <http://ideas.repec.org/a/ecm/emetrp/v56y1988i3p497-530.html>.
- Couch, K. A. (1992). New evidence on the long-term effects of employment training programs. *Journal of Labor Economics*, 10(4):380–88. available at <http://ideas.repec.org/a/ucp/jlabec/v10y1992i4p380-88.html>.
- Cox, R. (1972). Regression models and life tables. *Journal of the Royal Statistical Society*, 34(B):187–220.



- Danziger, S. H., Sandefur, G. D., & Weinberg, D. H. (1994). *Confronting poverty: prescriptions for change*. Harvard University, Cambridge, MA.
- Eberwein, C., Ham, J. C., & LaLonde, R. J. (1997). The impact of being offered and receiving classroom training on the employment histories of disadvantaged women: Evidence from experimental data. *Review of Economic Studies*, 64(4):655–82. available at <http://ideas.repec.org/a/bla/restud/v64y1997i4p655-82.html>.
- Fernandes, R., Menezes-Filho, N. A., & Zylberstajn, H. (2000). Avaliando o planfor: o programa de treinamento do sindicato dos metalúrgicos de são paulo. *Pesquisa e Planejamento Econômico*, 30(2).
- Fernandes, R. & Pazello, E. T. (2001). Avaliação de políticas sociais: incentivos adversos, focalização e impacto. In Lisboa, M. & Menezes-Filho, N., editors, *Microeconomia e sociedade no Brasil*, pages 151–171.
- Friedlander, D., Greenberg, D. H., & Robins, P. K. (1997). Evaluating government training programs for the economically disadvantaged. *Journal of Economic Literature*, 35(4):1809–1855. available at <http://ideas.repec.org/a/aea/jecclit/v35y1997i4p1809-1855.html>.
- Greene, W. H. (1998). *Econometric analysis*. Prentice-Hall, New Jersey, 3 edition.
- Ham, J. C. & LaLonde, R. J. (1996). The effect of sample selection and initial conditions in duration models: Evidence from experimental data on training. *Econometrica*, 64(1):175–205. available at <http://ideas.repec.org/a/ecm/emetrp/v64y1996i1p175-205.html>.
- Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1):153–61. available at <http://ideas.repec.org/a/ecm/emetrp/v47y1979i1p153-61.html>.
- Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *Review of Economic Studies*, 64(4):605–54. available at <http://ideas.repec.org/a/bla/restud/v64y1997i4p605-54.html>.
- Heckman, J. J., Lalonde, R. J., & Smith, J. A. (1999). The economics and econometrics of active labor market programs. In Ashenfelter, O. & Card, D., editors, *Handbook of Labor Economics*, volume 3 of *Handbook of Labor Economics*, chapter 31, pages 1865–2097. Elsevier. available at <http://ideas.repec.org/h/eee/labchp/3-31.html>.
- Heckman, J. J. & Smith, J. (1997). Making the most out of programme evaluations and social experiments: Accounting for heterogeneity in programme impacts. *Review of Economic Studies*, 64(4):487–535. available at <http://ideas.repec.org/a/bla/restud/v64y1997i4p487-535.html>.
- Heckman, J. J. & Smith, J. A. (1995). Assessing the case for social experiments. *Journal of Economic Perspectives*, 9(2):85–110. available at <http://ideas.repec.org/a/aea/jecper/v9y1995i2p85-110.html>.
- Hosmer JR., D. & Lemeshow, S. (1999). *Applied survival analysis*. John Wiley, New York.
- Kiefer, N. M. (1988). Economic duration data and hazard functions. *Journal of Economic Literature*, 26(2):646–79. available at <http://ideas.repec.org/a/aea/jecclit/v26y1988i2p646-79.html>.
- Kleinbaum, D. G. (1996). *Survival analysis: a self-learning text*. Springer-Verlag, New York.
- LaLonde, R. J. (1986). Evaluating the econometric evaluations of training programs with experimental data. *American Economic Review*, 76(4):604–20. available at <http://ideas.repec.org/a/aea/aecrev/v76y1986i4p604-20.html>.

- LaLonde, R. J. (1995). The promise of public sector-sponsored training programs. *Journal of Economic Perspectives*, 9(2):149–68. available at <http://ideas.repec.org/a/aea/jecper/v9y1995i2p149-68.html>.
- Manski, C. F. (1995). *Identification problems in the social sciences*. Harvard University, Cambridge, M.A.
- Menezes-Filho, N. A. & Picchetti, P. (2000). Os determinantes da duração do desemprego em são paulo. *Pesquisa e Planejamento Econômico*, 30(1):23–47.
- Menezes-Filho, N. A. & Picchetti, P. (2001). Desemprego. In Lisboa, M. & Menezes-Filho, N., editors, *Microeconomia e sociedade no Brasil*, pages 227–249. available at <http://ideas.repec.org/h/eee/labchp/3-31.html>.
- Rios-Neto, E. (2001). *Continuidade à metodologia de acompanhamento e avaliação do PLANFOR*. UFMG/CEDEPLAR, Belo Horizonte.
- Rios-Neto, E. & Assunção, R. (1998). *Metodologia de acompanhamento e avaliação do Plano Estadual de Qualificação (PEQ)*. UFMG/CEDEPLAR, Belo Horizonte.
- Rios-Neto, E. & Barros, R. P. (2000). *Uma proposta de avaliação do Plano Nacional de Formação Profissional (PLANFOR)*. UFMG/CEDEPLAR, Belo Horizonte.
- Rios-Neto, E. & Oliveira, A. (1999). Políticas voltadas para a pobreza: o caso da formação profissional. In Henriques, R., editor, *Desigualdade e pobreza no Brasil*. IPEA, Rio de Janeiro.
- Schiller, B. (1998). *The economics of poverty and discrimination*. Prentice Hall, New York.
- Schmid, G., O'Reilly, J., & Schömann, K. (1996). *International handbook of labour market policy and evaluation*. Edward Elgar, Aldershot, U.K.

A. ORGANIZAÇÃO DOS DADOS

A pesquisa gerou informações retrospectivas ao nível individual sobre a ocupação em cada mês e assim as durações, e a forma do banco de dados é uma história de eventos na qual o tempo de ocorrência da transição é então registrado. A seqüência de dados binários, igual a 1 se o indivíduo está ocupado no tempo t e 0 caso contrário, representa a história de eventos para cada indivíduo. Apesar da estrutura básica de dados parecer simples, é árduo colocar os dados nesta forma em estudos de histórias de vida que contêm informação sobre eventos repetidos. Se o estado de destino fosse absorvente, modelos de episódio único poderiam ser aplicados. Ao lidar com durações que representam intervalos de tempo para os quais nem todos os estados são absorventes, é possível que cada indivíduo apresente mais de uma transição. Neste estudo, nos interessa a duração dos episódios de desemprego, que podem ser repetidos para cada indivíduo. Uma abordagem seria conduzir uma análise separada para cada evento sucessivo, estimando um modelo para cada episódio, se assumimos que o modelo é diferente de um evento para o outro. Por outro lado, sendo o processo essencialmente o mesmo entre os eventos sucessivos, fazer uma análise separada para cada evento é estatisticamente ineficiente. A abordagem apropriada então é tratar cada intervalo entre eventos para cada indivíduo como uma observação separada. Assim, partimos de um banco de dados individuais, com informações mensais sobre o estado na força de trabalho e variáveis explicativas de interesse, como idade, grupo do experimento (tratamento=1 e controle=0), raça, etc., e o transformamos em um arquivo de dados orientado para eventos ou com episódios separados, organizando as observações como pseudo-observações. O banco de dados orientado para episódios permite que os indivíduos contribuam números variados de vezes à expressão de verossimilhança, dependendo de quando e se mudaram de estado ao longo do curso do painel. Além disto, o banco de dados orientado para episódios permite que os tempos ordenados de ocorrência do evento sejam agrupados entre os indivíduos para o cálculo das funções de sobrevivência e risco agregadas.



B. MÉDIAS DE VARIÁVEIS, SEGUNDO O GRUPO DO EXPERIMENTO, AMOSTRA TOTAL E DESEMPREGADOS

Tabela 7 – Médias de Variáveis, segundo o Grupo do Experimento, Amostra Total e Desempregados

	Total da Amostra		Desempregados	
	Tratamento	Controle	Tratamento	Controle
OCPDEZ96	0,38	0,51		
SEXO	0,82	0,7	0,77	0,64
BRANCA	0,29	0,31	0,31	0,31
JOVEM	0,54	0,42	0,66	0,62
CASADO	0,21	0,28	0,13	0,2
TEMFILHO	0,22	0,28	0,15	0,17
ESTUDA	0,51	0,4	0,62	0,59
SERIE4	0,3	0,22	0,33	0,21

C. MODELOS PARAMÉTRICOS

As variáveis utilizadas são: *trata* indica o grupo do experimento (tratamento=1 vs. controle=0); *ocpdez96* representa o estado na força de trabalho no momento do treinamento (1=ocupado vs. 0=desempregado); *sexo* (1=homens vs. 0=mulheres); *branca* (1=brancos vs. 0=outras raças); *jovem* (14-19 anos=1 vs. 20 anos ou +=0); *casado* (1=casado vs. 0=outros estados conjugais); *temfilho* (1=tem filhos vs. 0=não tem filhos); *estuda* (1=freqüente escola vs. 0=não freqüente escola); *serie4* (1=até 4 anos de estudo vs. 0=mais de 4 anos de estudo).

A tabela 8 mostra os resultados referentes ao teste log-rank.¹⁵ Para o total da amostra, o efeito da variável de tratamento é significativo, indicando que a participação no treinamento tem uma associação significativa com o tempo de sobrevivência no desemprego, o que vem a confirmar os resultados descritivos anteriores. O sinal negativo da estatística para o grupo do experimento indica menores durações para o grupo de tratamento, o que também reafirma o visto anteriormente.

Para a subamostra dos desempregados, a tabela mostra que o efeito da variável de tratamento é menos significativo, sugerindo a associação da participação no treinamento com o tempo de sobrevivência no desemprego, ainda que mais fraca do que para o total da amostra. O sinal negativo da estatística para o grupo do experimento também indica menores durações para o grupo de tratamento.

Testes por efeitos incrementais de adicionar variáveis ao conjunto foram feitos através de um procedimento de inclusão para frente, que consiste em primeiro encontrar a variável com a maior estatística χ^2 e colocá-la na série a ser testada, e depois encontrar a variável que produz o maior incremento no χ^2 conjunto das duas variáveis. Este χ^2 conjunto testa a hipótese nula de que os coeficientes das variáveis no modelo são ambas 0. O incremento do χ^2 é um teste da hipótese nula de que o coeficiente da segunda variável é 0 quando a primeira é controlada. Este processo é repetido até que todas as variáveis tenham sido incluídas. Para cada variável, é obtido um teste da hipótese de que a variável não tem efeito sobre o tempo de sobrevivência, controlando por todas as variáveis acima dela. O procedimento de inclusão pode levar a conclusões bastante diferentes da análise univariada; um dos motivos é que variáveis que antes tinham efeitos significativos podem ser relacionadas a outras covariáveis, que tenham ainda maiores efeitos sobre o tempo de sobrevivência. Isto é exatamente o que ocorre com a variável de tratamento entre os desempregados, cujos efeitos sobre a duração do desemprego perdem

¹⁵Os resultados referentes ao teste Wilcoxon são bastante similares, não diferindo na significância e sinal das variáveis.

Tabela 8 – Direção das Estatísticas de Teste e Valores e Significância dos χ^2 Univariados para o Teste Log-Rank

	Total		Desempregados	
	Estatíst.	χ^2	Estatíst.	χ^2
OCPDEZ96	-	51,87*		
BRANCA	+	2,23	+	0,37
JOVEM	+	9,73*	+	1,67
CASADO	+	2,21	+	9*
TEMFILHO	+	3,42***	+	13,53*
ESTUDA	+	10,38*	+	2,71***
SERIE4	+	13,12*	+	10,27*
SEXO	-	117,4*	-	94,99*
TRATA	-	7*	-	4,31**

* significante ao nível de 1%;

** significante ao nível de 5%;

*** significante ao nível de 10%;

Tabela 9 – Seqüência de Inclusão das Variáveis e Significância para o Teste Log-Rank

	Total da Amostra		Desempregados	
	Ordem	Signif.	Ordem	Signif.
SEXO	1	*	1	*
OCPDEZ96	2	*		
SERIE4	3	*	2	*
TRATA	4	*	5	
JOVEM	5	**	3	***
TEMFILHO	6	**	4	*
BRANCA	7		7	
ESTUDA	8		6	
CASADO	9		8	

* significante ao nível de 1%; ** significante ao nível de 5%; *** significante ao nível de 10%;

significância ao se controlar pelo sexo, nível de escolaridade, idade e presença de filhos. Para a amostra total, as conclusões são as mesmas da análise univariada.

Embora seja um procedimento útil para análises preliminares de dados de sobrevivência e para testar hipóteses simples sobre diferenças da sobrevivência entre grupos, a análise não paramétrica não permite examinar os efeitos das covariáveis. Por não fornecer estimativas dos coeficientes, não há como quantificar o efeito de uma covariável sobre o tempo de sobrevivência. Além disto, os testes incrementais não testam realmente o efeito de cada variável controlando por todas as outras, mas efeito de cada variável controlando pelas variáveis que já foram incluídas. Dado que não se tem controle sobre a ordem de inclusão, estes testes podem levar a conclusões erradas; entretanto, o método nos forneceu evidências iniciais importantes antes de procedermos à estimação de modelos de regressão.



D. ANÁLISE PARAMÉTRICA

Mesmo tendo uma idéia da forma da curva de risco através da análise não paramétrica e semi-paramétrica, estimamos alguns modelos paramétricos alternativos para buscar uma descrição mais refinada da sobrevivência dos grupos do experimento. Modelos paramétricos permitem introduzir o efeito de um grande número de características sobre a duração da permanência em um dado estado, sendo possível testar hipóteses sobre a forma da função de risco.

Em primeiro lugar, estimamos modelos de tempo de vida acelerado (*Accelerated Failure Time – AFT*), com o objetivo de verificar a possibilidade de escolher uma forma paramétrica que descreva a sobrevivência dos grupos estudados. Os modelos AFT fazem parte da classe de modelos de regressão que, em sua forma geral, descrevem uma relação entre as funções de sobrevivência de quaisquer dois indivíduos. Sendo $S_i(t)$ a função de sobrevivência para o indivíduo i , então para qualquer outro indivíduo j , o modelo AFT diz que: $S_i(t) = S_j(\phi_{ij}t)$ para todos os t , onde ϕ_{ij} é uma constante específica ao par (i, j) (Alisson, 1995). O que realmente estimamos é um caso especial deste modelo: sendo T_i uma variável aleatória que denota o tempo do evento para o indivíduo j na amostra e sendo x_{j1}, \dots, x_{jk} os valores das k covariáveis para o mesmo indivíduo, o modelo é:

$$\log T_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik} + \sigma \epsilon_i$$

onde ϵ_i é um termo de erro aleatório, e β_0, \dots, β_k e σ são parâmetros a serem estimados. Exponenciando a equação:

$$T_i = \exp\{\beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik} + \sigma \epsilon_i\}$$

As diferenças entre este modelo e o de regressão linear são que há um σ antes do ϵ , o que significa que a variância de ϵ é fixada em algum valor e σ muda de valor para acomodar mudanças na variância do erro; e a variável dependente é uma transformação logarítmica (de T), cujo propósito é assegurar que os valores preditos de T sejam positivos, independente dos valores dos x 's e β 's. Em um modelo de regressão linear, assume-se que ϵ_i tem uma distribuição normal com média e variância constantes, e que os ϵ 's são independentes entre as observações. O modelo AFT log-normal tem exatamente estes pressupostos. Os outros modelos AFT permitem distribuições de ϵ além da normal, mas retêm os pressupostos de média e variância constantes e de independência entre observações. Na análise de sobrevivência, é importante considerar estas distribuições alternativas de ϵ , porque elas têm diferentes implicações para as funções de risco que podem levar a diferentes interpretações substantivas. Algumas destas distribuições, para as quais há uma distribuição correspondente de T são: valores extremos (1 parâmetro) – exponencial, valor extremo (2 parâmetros) – Weibull, log-gamma – gamma e logística – log-logística.

Os modelos AFT são estimados pelo método de máxima verossimilhança (ML), que lida bem com dados que contém informações censuradas. O princípio básico de ML é escolher as estimativas cujos valores maximizam a probabilidade de observar o que de fato é observado. Primeiro é construída uma função de verossimilhança, que é uma expressão para a probabilidade dos dados como função dos parâmetros desconhecidos. Deve-se especificar um modelo dentre os descritos acima, o que significa escolher uma distribuição de probabilidade para a variável dependente e uma forma funcional que relacione os parâmetros desta distribuição aos valores das covariáveis. Em seguida, é feita a maximização, achando os valores dos parâmetros desconhecidos que tornam o valor desta expressão o maior possível. Isto requer um método numérico iterativo, ou seja, que envolve aproximações sucessivas.

Para a amostra total, a estatística χ^2 do multiplicador de Lagrange calculada para o modelo exponencial não rejeitou a hipótese nula de que o parâmetro de escala fosse igual a 1, o que sugere que a função de risco de saída do desemprego para a amostra total poderia ser constante ao longo do tempo. Quando estimado o modelo Weibull, o parâmetro de escala calculado foi maior que 1, indicando que o risco diminui com o tempo na seguinte proporção: um aumento de 1% no tempo desde o tratamento

Tabela 10 – Estimativa dos Coeficientes e Significância para a Amostra Total

	Exponencial	Weibull	Log-Normal	Log-Logístico	Gamma Gener.
Intercepto	2,8645*	2,8637*	2,198*	2,2154*	1,7995*
OCPDEZ96	-0,383*	-0,3832*	-0,2858*	-0,2924*	-0,2197*
BRANCA	0,0785	0,0782	0,0627	0,0308	0,0743
JOVEM	0,1874*	0,1878*	0,1941*	0,2178*	0,1739*
CASADO	0,0195	0,019	0,0179	-0,0088	0,0125
TEMFILHO	0,2192**	0,2194**	0,1288	0,1072	0,1045
ESTUDA	0,0512	0,0506	0,0409	0,034	0,0458
SERIE4	0,1846*	0,185*	0,1396*	0,1434*	0,1078**
SEXO	-0,7355*	-0,7386*	-0,6354*	-0,6637*	-0,5327*
TRATA	-0,2278*	-0,2285*	-0,1482**	-0,1535**	-0,0982
Escala (σ)	1	1,0163	1,0992	0,6483	1,0812
Forma	1	0,984			-0,5324
Log-Verossimilh.	-3538,7248	-3538,256	-3364,3293	-3401,1357	-3342,4558
Multipl.Lagrange	0,9147				

* significante ao nível de 1%;

** significante ao nível de 5%;

*** significante ao nível de 10%;

Número de Observações: 2427; Valores Censurados: 427.

Tabela 11 – Teste da Razão de Verossimilhança entre os Modelos, Amostra Total

1 g.l.	Exponencial x Weibull	0.9377	Aceita
2 g.l.	Exponencial x Gamma	3.925.381	Rejeita
1 g.l.	Weibull x Gamma	3.916.004	Rejeita
1 g.l.	LogNormal x Gamma	4.374.712	Rejeita

produziria uma diminuição de 1,6% no risco. O modelo log-normal estimado revelou um parâmetro de escala maior que 1, indicando que o risco apresenta um pico rapidamente, tal que a função é similar à Weibull e à log-logística. Quanto ao modelo log-logístico, a estimativa da escala foi menor que 1, implicando que a função estimada de risco segue a forma de U invertido; dada a similaridade dos riscos log-normal e log-logístico, os resultados são similares. No modelo gamma generalizado ajustado, os parâmetros de escala (1,08) e de forma (-0,53) estimados não permitiram inferir a adequação de nenhum caso especial, mas uma maior aproximação do modelo log-normal (parâmetro de forma = 0).

Todos os modelos produziram estimativas similares de coeficientes e p-valores, sendo necessária alguma maneira de decidir entre eles, já que esta escolha tem diferentes implicações para a forma da função de risco. Utilizamos então um método para escolher os melhores ajustes, através da comparação dos valores das log-verossimilhanças. Para a amostra total, dentre estes modelos, as menores magnitudes das log-verossimilhanças foram fornecidas pelos modelos gamma generalizado, log-normal e log-logístico. Além desta indicação, fizemos testes de qualidade do ajuste com a estatística da razão de verossimilhança para comparar modelos hierarquizados.¹⁶ Quando comparado ao Weibull, o modelo exponencial deveria ser aceito, implicando que o risco seria constante ao longo do intervalo. Isto é

¹⁶Um modelo é um caso especial de outro se pode ser obtido pela imposição de restrições sobre os parâmetros, o que é o caso dos modelos AFT, que englobam vários sub-modelos que diferem na distribuição pressuposta da duração.

**Tabela 12** – Estimativa dos Coeficientes e Significância para os Desempregados

	Exponencial	Weibull	Log-Normal	Log-Logístico	Gamma Gener.
Intercepto	2,7644*	2,7612*	2,1266*	2,131*	1,7889*
BRANCA	0,0483	0,0475	0,0506	0,0169	0,0731
JOVEM	0,2132*	0,215*	0,2286*	0,2476*	0,2122*
CASADO	0,0061	0,0051	0,0206	-0,0004	0,0175
TEMFILHO	0,3892*	0,3927*	0,2823**	0,2532***	0,2458**
ESTUDA	0,0664	0,0662	0,0557	0,0513	0,0542
SERIE4	0,197*	0,1974*	0,1413**	0,138**	0,1118***
SEXO	-0,7438*	-0,7524*	-0,6694*	-0,6979*	-0,5857*
TRATA	-0,1511***	-0,1513***	-0,0834	-0,0698	-0,0542
Escala (σ)	1	1,0383	1,1398	0,6757	1,1291
Forma	1	0,9631			-0,4957
Log-Verossimilh.	-2670,5842	-2668,7261	-2549,1544	-2580,7342	-2535,4439
Multipl.Lagrange	3,5137***				

* significante ao nível de 1%;

** significante ao nível de 5%;

*** significante ao nível de 10%;

Número de Observações: 1816; Valores Censurados: 343.

Tabela 13 – Teste da Razão de Verossimilhança entre os Modelos, Desempregados

1 g.l.	Exponencial x Weibull	37.161	Aceita
2 g.l.	Exponencial x Gamma	2.702.806	Rejeita
1 g.l.	Weibull x Gamma	2.665.645	Rejeita
1 g.l.	LogNormal x Gamma	274.211	Rejeita

consistente com o resultado anterior do teste do multiplicador de Lagrange; todavia, este modelo e os outros casos especiais do modelo gamma generalizado são rejeitados na comparação com este.

Concluindo, nossa escolha de modelo para a amostra total recaiu sobre os modelos log-logístico e gamma generalizado. O modelo log-logístico não se encaixa no esquema dos modelos hierarquizados, ou seja, não é um caso especial de outros e portanto não pôde ser a eles comparado pela razão de verossimilhança. Como apresentou uma menor log-verossimilhança do que os outros modelos e bastante próxima do modelo gamma generalizado, sugere um bom ajuste aos dados. Além disto, a desvantagem de se optar somente pelo modelo gamma generalizado é que sua função de risco é muito complicada, sendo difícil julgar a forma da função de risco a partir dos parâmetros estimados (Alisson, 1995).

A tabela 14 mostra a estimativa dos coeficientes das covariáveis para os modelos escolhidos para a amostra total – log-logístico e gamma generalizado. Os valores estimados dos coeficientes e sua significância são similares, portanto a análise aqui recai sobre o modelo log-logístico. Nesta métrica mostrada, suas magnitudes numéricas não são muito informativas, mas uma transformação simples leva a uma interpretação intuitiva: e^{β} fornece a razão estimada dos tempos de sobrevivência médios esperados para os grupos. Por exemplo, para a variável de tratamento, $e^{-0,1535} = 0,8577$ indica que, controlando pelas outras covariáveis, a duração dos episódios de desemprego para o grupo de tratamento é 14,2% menor do que para o grupo de controle.

Para os desempregados, os resultados da análise paramétrica são semelhantes aos da amostra total. Uma diferença foi a estatística χ^2 do multiplicador de Lagrange calculada para o modelo exponencial,

Tabela 14 – Estimativa dos Coeficientes e Significância para o Total da Amostra

	Log-Logístico	Gamma Gener.
Intercepto	2,2154*	1,7995*
OCPDEZ96	-0,2924*	-0,2197*
BRANCA	0,0308	0,0743
JOVEM	0,2178*	0,1739*
CASADO	-0,0088	0,0125
TEMFILHO	0,1072	0,1045
ESTUDA	0,034	0,0458
SERIE4	0,1434*	0,1078**
SEXO	-0,6637*	-0,5327*
TRATA	-0,1535**	-0,0982
Escala (σ)	0,6483	1,0812
Forma (δ)		-0,5324
Log-Verossimilh.	-3401,1357	-3342,4558

*significante ao nível de 1%;

**significante ao nível de 5%;

***significante ao nível de 10%;

Núm. de Observ.:2427;Valores Censurados:427.

que rejeitou a hipótese nula de que a função de risco de saída do desemprego para a amostra total seria constante ao longo do tempo. O parâmetro de escala estimado pelo modelo Weibull foi maior que 1, indicando que o risco diminui com o tempo, mas em um ritmo maior (3,8%). No modelo gamma generalizado, os parâmetros estimados também não permitiram inferir a adequação de nenhum caso especial. Novamente, os modelos produziram estimativas similares de coeficientes e p-valores, sendo necessária alguma maneira de decidir entre eles. Os modelos que melhor descrevem a função de risco verdadeira dos desempregados parecem ser gamma generalizado, log-normal e log-logístico. Os testes de qualidade do ajuste com a estatística da razão de verossimilhança para comparar modelos hierarquizados indicaram que, quando comparado ao Weibull, o modelo exponencial deveria ser aceito, o que, neste caso, não é consistente com o resultado anterior do teste do multiplicador de Lagrange. Todos os modelos foram rejeitados na comparação com o modelo gamma generalizado. Portanto, a escolha de forma do risco para os desempregados também recaiu sobre os modelos log-logístico e gamma generalizado.

Os coeficientes destes modelos para a subamostra de desempregados estão na tabela 15. Enfatizando os valores estimados pelo modelo log-logístico, a grande diferença em relação ao total da amostra é a perda de significância da variável indicativa do grupo do experimento. Isto significa que, controlando pelas outras covariáveis, a duração dos episódios de desemprego para o grupo de tratamento é estatisticamente equivalente à do grupo de controle.



Tabela 15 – Estimativa dos Coeficientes e Significância para os Desempregados

	Log-Logístico	Gamma Gener.
Intercepto	2,131*	1,7889*
BRANCA	0,0169	0,0731
JOVEM	0,2476*	0,2122*
CASADO	-0,0004	0,0175
TEMFILHO	0,2532***	0,2458**
ESTUDA	0,0513	0,0542
SERIE4	0,138**	0,1118***
SEXO	-0,6979*	-0,5857*
TRATA	-0,0698	-0,0542
Escala (σ)	0,6757	1,1291
Forma (δ)		-0,4957
Log-Verossimilh.	-2580,7342	-2535,4439

* significante ao nível de 1%;

** significante ao nível de 5%;

*** significante ao nível de 10%;

Núm. de Observ.:1816;Valores Censurados:343.